Video Game Recommendation Comparison

Name	Course
Ronen H	SI 671: Data Mining

The increasing prominence of video games over the past few years has led to a need for a robust video game recommendation system addressing a diverse community of users.

Introduction

Many people across the world enjoy playing video games for various reasons such as relieving stress, navigating through difficult missions, and going through intriguing plots. A question that arises is how to recommend video games to users with vastly different preferences. After all, a user might rely heavily on existing reviews or go against the grain.

Data

A common way to purchase video games in online through Amazon. A reviewer, verified or not, can leave behind a rating from 1 star (worst) to 5 stars (best) and can be voted in agreement by other users. Each video game has an associated set of categories. The data was collected via webscraping by Jianmo Ni of UCSD and spans from May 1996 to October 2018 (1).

Methods

The three video game recommendation systems to compare in terms of their benefits and downsides are:

1. **Based on similar categories.**
   - **The video game categories are represented as an itemset.**
   - **From the user video games, for each unique categories, obtain top 10 video games by Jaccard Similarity of at least 0.5 then average rating then popularity.**
2. **Based on rating, influence, and reliability of reviews.**
   - **The video games (nodes) and the reviews (edges) are represented as a graph. There is a directed edge from one video game to the other if a user has bought both. The weight is the modified average rating with a rating duplicated once if verified and then duplicated by the number of votes if any.**
   - **For each user video game, obtain top 10 video games from Louvain Community Detection by average rating then popularity. If no user video games in graph, defaulted to based on similar categories.**
3. **Based on similar users.**
   - **The user ratings are represented as a vector.**
   - **From the user ratings, find similar users with a Pearson Correlation Coefficient of at least 0.7 and obtain their liked (rating of at least 4) video games. If no similar users, defaulted to based on similar categories.**

After obtaining candidate video games, recommend top 10 video games based off popularity then average rating.

References

1. Ni, Jianmo, Jiacheng Li and Julian McAuley. *Empirical Methods in Natural Language Processing (EMNLP)*. 2019. https://cseweb.ucsd.edu//~jmcauley/pdfs/emnlp19a.pdf.

Analyses

The four users of interest are:

- User A: 2 reviews. Liked Sony PSP fighting. Disliked Nintendo DS platform.
- User B: 4 reviews. Interest in PC aerial combat games.
- User C: 5 reviews. Interest in Wii adventure games.
- User D: 8 reviews. Interest in Xbox 360 / PlayStation action role-playing games.

	title	jaccard_similarity
12105	PixelJunk Monsters [Online Game Code - Full G...	1.0
64498	Far Cry 3 - PS3 [Digital Code]	1.0
68132	Resident Evil Code Veronica X HD - PS3 [Digita...	1.0
61503	Thomas Was Alone (3 Way Cross Buy) - PS4 / PS3...	1.0
67335	Shin Megami Tensei: Persona 3 FES - PS3 [Digit...	1.0

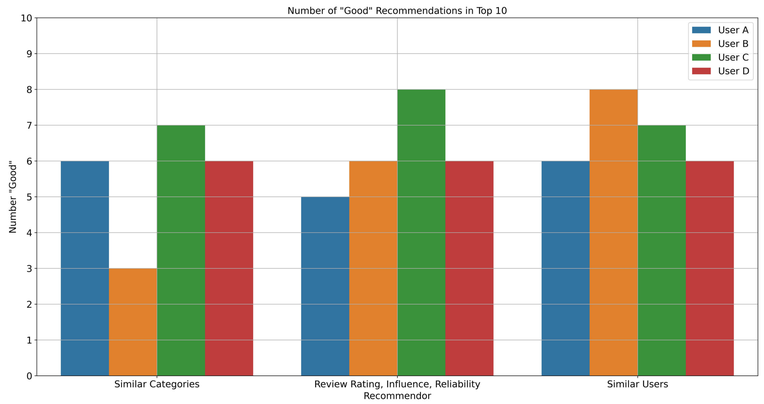
Based on similar categories, top 5 candidates from "Kingdoms of Amalur: Reckoning" for User D.

	title	average_rating
66399	Ryu ga Gotoku: Kiwami - PS4 [Japan Import]	5.0
32398	Kingdom Hearts: Chain of Memories	5.0
31847	Crazy Taxi	5.0
33227	Future Cop: L.A.P.D.	5.0
53715	Metro 74" TV Stand Finish: Black Woodgrain	5.0

Based on review rating, influence, and reliability, top 5 candidates from "Carnival Games" for User C.

	user	number_reviews	pearson_coefficient
19219	A13251HBWRMQ8W	15	1.0
12091	A1A87TUBBDPMF0	7	1.0
93986	A3KB431ELEHOCH	6	1.0
75131	A4OFH715NQNB1	4	1.0
45557	A1K0NTOHXNIWX6	3	1.0

Based on similar users, similar users to User B to obtain candidates from.



Bar plot of number of "good" (sensible) reviews in the top 10 based off user reviews.

Results

After observing the themes and "goodness" (sensibility) of the top 10 recommendations, the benefits and downsides of each recommender system are

1. **Based on similar categories:**
   - **Benefit: Good for users with barely any purchases like User A. Explore other interests within preferred video game categories.**
   - **Downside: Bad for users with uncommon interests like User B. Interest not well captured by video game categories.**
2. **Based on rating, influence, and reliability of reviews:**
   - **Benefit: Good for users with well-defined interests like User C. Adds diversity with a mix of games and accessories.**
   - **Downside: Bad for users with barely any purchases like User A. Does not have a notion of potential interests.**
3. **Based on similar users:**
   - **Benefit: Good for users with uncommon interests like User B. Good for users with a good number of purchases like User D. Incorporates topical similarity.**
   - **Downside: Bad for users with barely any purchases like User A. Unable to find similar users (defaulting to based on similar categories).**